Journal of Computers Vol. 33 No. 5, October 2022, pp. 109-119
doi: 10.53106/199115992022103305010

A Super-resolution Reconstruction Method for Underwater Fish Images with Feedback Networks

Xiu-Juan Wang[1,3], Liang Mao[2*], Ji Wang[1,3], Wen-Chao Gong[3]

[1] College of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang, Guangdong, 524088, China
wxj5065@163.com, wangji@gdou.edu.cn
[2] Guangdong-Hong Kong-Macao Greater Bay Area Artificial Intelligence Application Technology Research Institute, Shenzhen, Guangdong, 518055, China
maoliangscau@163.com
[3] Research Center of Guangdong Smart Oceans Sensor Networks and Equipment Engineering, Zhanjiang, Guangdong, 524088, China
wangji@gdou.edu.cn, 964971895@qq.com

Received 9 December 2021; Revised 3 February 2022; Accepted 3 March 2022

Abstract. Video surveillance is commonly used for production process monitoring in aquaculture. Because underwater video images are prone to blurring and unclear detail texture, this paper proposes a super-resolution network reconstruction method with a feedback network. The reconstruction effect is enhanced by a feedback network in the process of mapping low-resolution images to high-resolution images, which is implemented with a constrained recurrent neural network (RNN) to process deep feedback information. However, the shallow features are fed to the feedback net-work module to generate deeper features through multiple upsampling and downsampling and progressive refinement, which give the network structure an early reconstruction capability and are conducive to more realistically reconstructing high-resolution images. A course learning strategy is introduced to make the network applicable to more complex tasks and improve its robust-ness. We construct a training set of 800 images and validation and test sets of 10 images each using images of marine net fish as the research object. The method of this paper is validated on the public dataset Set14 and a self-built dataset. The proposed method outperforms other methods in both subjective and objective evaluations on the public and self-built datasets. Good foundation for high-definition monitoring of aquaculture is laid.

Keywords: aquaculture, fish images, super-resolution reconstruction, feedback networks, course learning

1 Introduction and Related Works

The Litchi With scientific and technological development, the intelligent fishery is in-evitable. The monitoring of aquaculture objects is an important task. However, under-water video surveillance generally has the problems of blurred images and low clarity, which leads to the inability to detect abnormal objects in a timely manner. If aquacultured fish suffering from infectious diseases are not salvaged in time, they will cause water pollution, causing more farmed objects to become ill, resulting in economic loss. [1-2] Therefore, to obtain clearer video image data, super-resolution reconstruction [3] is used to transform images from low-definition to high-resolution. Super-resolution reconstruction is widely used, such as in high-definition video surveillance, video restoration, and medical imaging. [4] Traditional super-resolution methods are categorized as interpolation-, reconstruction-, and learning-based. [5] While these methods are effective, they suffer from the problems of difficult feature extraction and high computational effort.

With the rapid development of deep learning, researchers have incorporated convolutional neural networks (CNNs) [6] and generative adversarial networks (GANs) [7] in super-resolution processing and made corresponding improvements to obtain net-works with super-segmentation properties. [8] SRCNN, a super-resolution deep learning method for single images proposed by Chao et al., [9] performs multiple convolution processes on low-resolution images to obtain super-resolution images; Chao et al. [10] improved the speed of training models in SRCNN to realize FSRCNN; and Shi et al. [11] proposed ESPCN, whose sub-pixel convolution layer extracts features directly from low-resolution images and rearranges them to obtain high-resolution images, which effectively reduces computational complexity. Kim et al. [12-13] proposed VDSR, which incorporates a residual network [14] to accelerate the convergence of network training, and DRCN, which incorporates residual and recursive neural net-work ideas and uses recurrent neural networks (RNNs) for nonlinear mapping to effectively reduce network complexity. Ledig et al. [15] first used GANs on image super-resolution reconstruction studies and

* Corresponding Author

proposed SRGAN, which uses GANs to complement the details in generated images and increase their realism. Tong et al. [16] pro-posed SRDenseNet, which uses dense connectivity to connect the features of layers in series to improve performance. Kuang [17] proposed the use of special information learning to improve super-resolution image reconstruction. The above methods use a feedforward network structure, i.e., a later layer can obtain information from a former layer but not vice versa. However, while this improves image resolution and some-what reduces computational complexity, the number of parameters increases as the network deepens, which occupies much memory, increases the computational burden, and tends to overfit during training.

A super-resolution method with a feedback network is necessary to solve the problems of the feedforward network structure. Drawing on the successful applications of feedback networks such as in target recognition, medical image segmentation, and medical image super-resolution, Zhou et al. [18] added a feedback network to a network to detect fatigued driving, achieving self-evolution and thus improving the adaptability to a user's personalized fatigue characteristics. An et al. [19] built a feed-back module in the image feature extraction layer to extract more effective features, solving the difficulty of segmenting medical images and extracting effective features. Chen et al. [20] proposed a feedback adaptive weighted dense network method based on a trusted depth CNN, in which a feedback connection passes the information of the out-put image to the lower-layer features for a clearer reconstructed image. The method is mainly used to reconstruct medical images to improve their resolution and has not been applied to underwater image processing.

We propose a super-resolution reconstruction method based on CNNs, incorporating feedback networks and course learning strategies implemented by constrained RNNs containing hidden states [21-22] for underwater video images with textures, un-clear details, and deep feedforward network structures with many parameters, using seawater net-culture fish captured by the GLADIUS MINI robotic whale as the re-search object. A top-down feedback information flow provides deep information to re-fine the shallow information and progressively generate high-resolution images. The feedback network uses G-groups of dense jump-connected upper and lower sampling layers to progressively generate robust deep information representations, and the hidden output information of each iteration is transmitted to the next iteration to modulate the input. The method extracts the shallow information of a low-resolution image, feeds it into a feedback network to gradually optimize it, and fuses the upsampled information of the low-resolution image with the generated deep information features to output a high-resolution image. A network model to reconstruct a high-resolution image using low-resolution images is constructed using an end-to-end training approach.

2 Super-resolution Reconfiguration of Network Structure

The proposed super-resolution reconstruction network incorporates a constrained RNN implementation containing hidden states on the basis of a CNN. A feedback network (FB) processes the feedback information stream and generates a powerful deep representation, refining the shallow information with the deep information so that the reconstructed image can restore more image details. Fig. 1 shows the network structure, which can be expanded to T iterations, where T is set to 4. Each iteration can be divided into a low-resolution feature extraction module, feedback network module, and reconstruction module. The low-resolution feature extraction module consists of Conv (3, m) and Conv (3, 4 m) (m is the number of convolution kernels) to extract the shallow information features F_{in}^{t} of the low-resolution image for input to the feedback network module, which obtains the deep information by upsampling and downsampling, and its output information F_{in}^{t} is input information to the feedback network module. The output information F_{out}^{t} is passed to the feedback network module in the (t+1)th iteration and the reconstruction module, which is used to reconstruct the final high-resolution image, using the deconvolution layer Deconv to recover the output information F_{out}^{t} passed by the last feedback network module to the high-resolution image, using Conv (3, n) to generate the residual image I_{Res}^{t}, and the output image I_{SR}^{t} is calculated as

$$I_{SR}^{t} = I_{Res}^{t} + f_{up}(I_{LR}) . \tag{1}$$

where f_{up} indicates upsampling, which in this paper is bilinear. The upsampled images are delivered using a global residual jump connection. The bilinear upsampling is calculated as

$$f(x,y)=\frac{f(Q_{11})}{(x_2-x_1)(y_2-y_1)}(x_2-x)(y_2-y)+\frac{f(Q_{21})}{(x_2-x_1)(y_2-y_1)}(x-x_1)(y_2-y)+\frac{f(Q_{12})}{(x_2-x_1)(y_2-y_1)}(x_2-x)(y-y_1)+\frac{f(Q_{22})}{(x_2-x_1)(y_2-y_1)}(x-x_1)(y-y_1). \tag{2}$$

where $f(x,y)$ is the requested pixel point; $f(Q_{11})$, $f(Q_{12})$, $f(Q_{21})$, and $f(Q_{22})$ are the pixel points to the lower-left, upper-left, lower-right, and upper-right, respectively, of the f point in the direction of the coordinate system; and (x_1,y_1), (x_1,y_2),(x_2,y_1) , and (x_2,y_2) are the corresponding values of the pixel points.

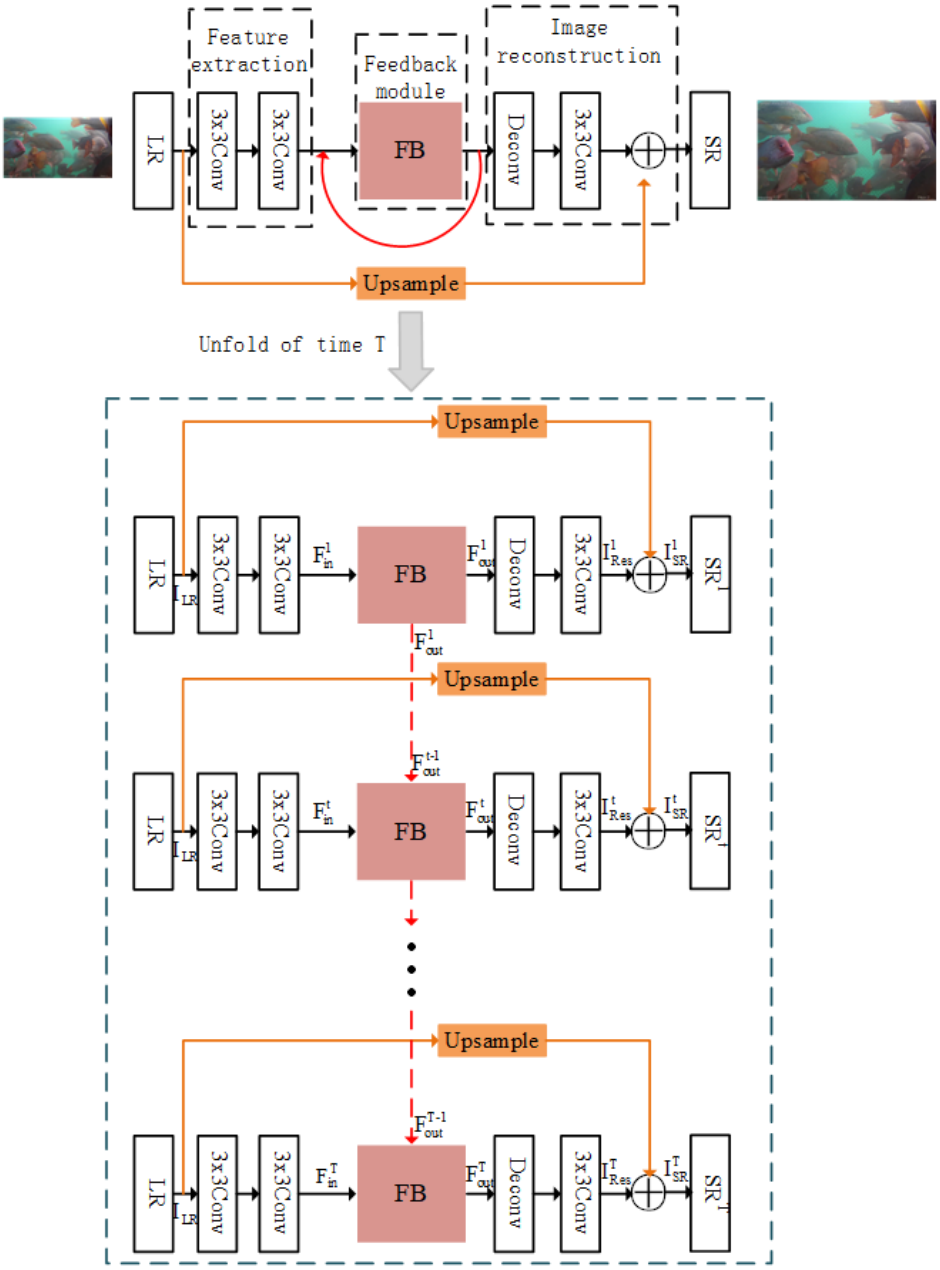

Fig. 1. Super-resolution network architecture

Remark. The FB module is the improvement of this paper. The green frame at the bottom is the T-unfold structure. LR and SR represent input and output images, respectively. Conv is the convolution layer, and 1 x 1 and 3 x 3 are the sizes of the convolution kernels. Deconv is the deconvolu-tion layer. Upsample represents the upsampling operation of low-resolution images, which is transmitted in the form of a global-intensive skip connection.

2.1 Feedback Network Module

The feedback network structure in the proposed method is shown in Fig. 2, from which it can be seen to consist of network sub-structures mapped by G groups of features, where G is set to 3. The sub-structures adopt a dense jump connection, i.e., the output of the front Conv is added to the input of the back Conv, and the output of the front Deconv is added to the input of the back Deconv to prevent gradient disappearance. Each sub-network completes the mapping between high- and low-resolution images by upsampling and downsampling. Taking iteration t as an example, the output information F_{out}^{t-1} and shallow information F_{in}^{t} of the (t-1)th feedback module are used as input information, the output information F_{out}^{t-1} is used to correct the shallow information F_{in}^{t}, and the output information F_{out}^{t} is used as input for the (t+1)th iteration. The information of F_{out}^{t-1} and F_{in}^{t}, stitched together in the network, is compressed using Conv (1, m), the input feature L_0^t is obtained by refining the information of F_{in}^{t}, the low-resolution feature L_g^t is compressed by the feature mapping substructure Conv (1, m) of the gth group for $([H_1^t, H_2^t, \cdots, H_g^t])$, and the high-resolution feature H_g^t is obtained by the feature mapping sub-structure Deconv (k, m) of group g (k denotes the convolution kernel size) by upsampling on $([L_0^t, L_1^t, \cdots, L_{g-1}^t])$. A 1×1 Conv is used between upsampling and downsampling to reduce the computational effort. The output information F_{out}^{t} is obtained by Conv (1, m) for efficient fusing of the low-resolution feature information $([L_1^t, L_2^t, \cdots, L_G^t])$.

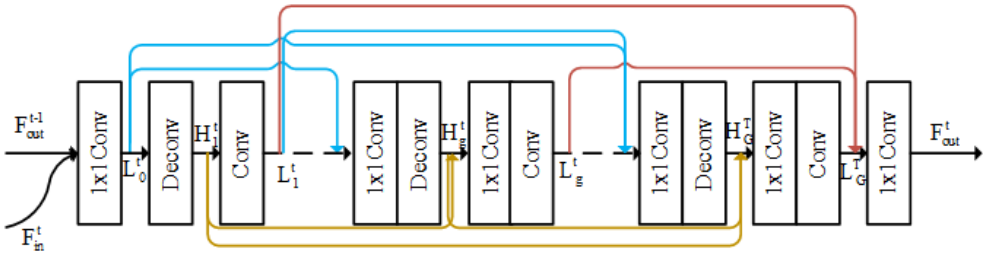

Fig. 2. Feedback network module structure

2.2 Course Learning Strategy

The course learning strategy gradually increases the difficulty of the learning target from simple to complex and can be used to improve the training process. We use it to gradually recover corrupted low-resolution images and reconstruct more realistic high-resolution images. Its network model is unfolded into T iterations, obtaining T target high-resolution images each time. Course learning ranks them by training difficulty, the loss of each iteration is counted as part of the total loss, and the L1 loss function [23] is used for optimization, as follows:

$$L(\theta) = \frac{1}{T}\sum_{t=1}^{T} W^t \, || \, I_{HR}^T - I_{SR}^T \, ||_1 . \tag{3}$$

Where W^t is the output at the iteration t, which is a constant; θ is the set of network parameters; T is the total number of iterations; I_{HR}^T denotes the target high-resolution image; and I_{SR}^T denotes the high-resolution image generated by the super-resolution network.

3 Experimental Dataset and Model Training

3.1 Feedback Self-built Dataset

The self-built dataset was collected using the GLADIUS MINI underwater ma-chine whale (Shenzhen Diving Innovation Technology Co., Ltd.), as shown in Fig. 3. The project team recorded fish videos on November 11, 2020, a sunny day with a temperature of about 25°C, in the seawater net box culture area of Nan'ao District,

Shenzhen City, Guangdong Province, China. Each net box was 6 meters long, 4 meters wide, and 5 meters deep. Images of the filmed videos were manually extracted chronologically, and 800 images were randomly extracted as the training set, 10 as the validation set, and 10 as the test set. Fig. 4 shows high- and low-resolution images at different magnifications and locally magnified images. Fig. 4(a) shows a high-resolution image, and Fig. 4(b) to Fig. 4(d) are low-resolution images at magnifications of 2x, 3x, and 4x, respectively. Fig. 4(e) to Fig. 4(h) show different magnifications of the information in the red boxes in Fig. 4(a) to Fig. 4(d). It can be seen that the details of the low-resolution images are blurrier.

Fig. 3. GLADIUS MINI whale

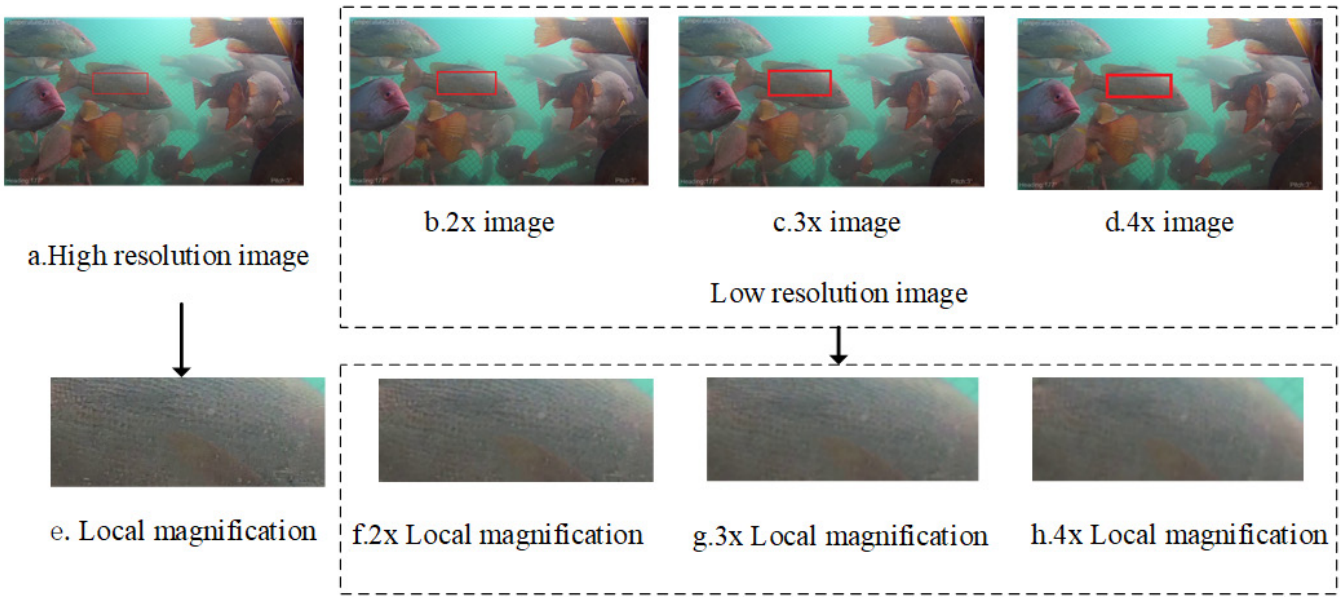

Fig. 4. High-and low-resolution images at different magnifications; local magnification images

3.2 Experimental Platform

The experimental hardware platform was an Intel Core i7-9700 CPU with 16 GB of RAM and a GeForce RTX 2070 GPU, and the software platform was a Windows 10 operating system, using the Python 3.7 programming language and PyTorch 1.1 deep learning framework.

3.3 Training Strategies

The proposed method adopts an end-to-end approach, the ReLU activation function is chosen for training, and the Adam optimizer is used. Images with magnifications of 2x, 3x, and 4x are trained separately. The parameters are set as follows. The batch sizes are 64, 32, and 16 for 2x, 3x, and 4x images, respectively; the initial learning rate is 1e-5; there are 200 epochs; the learning rate changes to 0.1 times the previous one at epochs 40, 60, and 150; and the training model is saved once every 50 epochs. Because the image clarity varies under different magnifi-

cations, some parameters are inconsistent in the training process for images with different magnifications in order to get a better reconstruction effect, specifically: When training images with magnification of 2x, the convolution kernel size of convolution layer and deconvolution is 6×6, and the step stride and padding size are set to 2. When training images with a magnification of 3x, the size of the convolution kernel of the convolution layer and deconvolution is 7×7, the step stride is 3, and the padding size is 2. When the training magnification is 4x, the size of the convolution kernel of the convolution layer and deconvolution is 8×8, the step stride is 4, and the padding size is 2.

4 Experimental Results and Analysis

4.1 Common Evaluation Indicators

In the evaluation of hyper-segmentation image effects, [24] professionals are required to judge the image quality, which is heavily subjective, making a unified evaluation standard difficult. Image quality is judged by metrics [25-26] such as the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), whose higher values in-dicate less image distortion and higher image quality.

PSNR is calculated as

$$\mathrm{PSNR}=10*\log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right). \tag{4}$$

where MAX is the maximum pixel value, which we take as 255, and MSE is the mean squared difference between the corresponding pixels of images,

$$\mathrm{MSE}=\frac{1}{M*N}\sum_{i=1}^{N}\sum_{j=1}^{M}(f_{ij}-f'_{ij})^2. \tag{5}$$

where M and N are the length and width, respectively, of an image; and f_{ij}, f'_{ij} are the respective pixel values of the high- and low-quality image.

SSIM is calculated as

$$SSIM=L(x,y)*C(x,y)*S(x,y). \tag{6}$$

where $L(x,y)$, $C(x,y)$, and $S(x,y)$ are the comparisons of brightness, contrast, and structure, respectively, between images, calculated as

$$L(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1}. \tag{7}$$

$$C(x,y)=\frac{\sigma_x\sigma_y+C_2}{\sigma_x^2\sigma_y^2+C_2}. \tag{8}$$

$$S(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3}. \tag{9}$$

where μ_x and μ_y are the mean pixel values of images x and y, respectively; σ_x and σ_y are the respective standard values of their pixels; $\sigma_x\sigma_y$ is the covariance of images x and y; and C_1, C_2, C_3 are constants to avoid zero denominators.

4.2 Results on Public Dataset

The proposed method was trained with the bicubic method [27], FSRCNN, and SRGAN on the common dataset DIV2K. Fig. 5 shows the loss value change curves obtained by the proposed method for training on images with different magnifications, where red, blue, and green lines indicate curve changes when 2x, 3x, and 4x images, respectively, are trained. It can be seen that 2x and 3x images can converge quickly when trained, fluctuating more in the first 70 epochs when 4x images are trained and gradually tending to stabilize. The final loss values of the three magnifications of images from 175 to 200 epochs remain around 0.2, and the model converges well.

The best training network model was tested on Set14, as shown in Fig. 6 and Table 1. From Fig. 6, it can be seen that the images reconstructed by the bicubic method are visually blurred, the reconstructed results of FSRCNN have fine textures that are not in the original images, and the reconstructed images of SRGAN have deviations in color from the original images and generate more severe textures. The reconstructed image from the proposed method is closer to the original, as can be seen in Table 1. The PSNR and SSIM of the proposed method reached 28.40 and 0.8432, respectively, at 2x magnification; 29.88 and 0.8329 at 3x; and 28.40 and 0.7769 at 4x. The PSNR and SSIM of the bicubic method reached 28.08 and 0.8087, respectively, at 2x; 27.89 and 0.7953 at 3x; and 27.17 and 0.7180 at 4x. The PSNR and SSIM of FSRCNN reached 29.09 and 0.8203, respectively, at 2x; 28.03 and 0.8044 at 3x; and 27.62 and 0.7544 at 4x. The PSNR and SSIM of SRGAN reached 24.29 and 0.8354, respectively, at 2x; 26.17 and 0.7691 at 3x; and 23.08 and 0.6901 at 4x. In summary, the evaluation indexes of the proposed method were higher than those of the other methods for images with different magnifications.

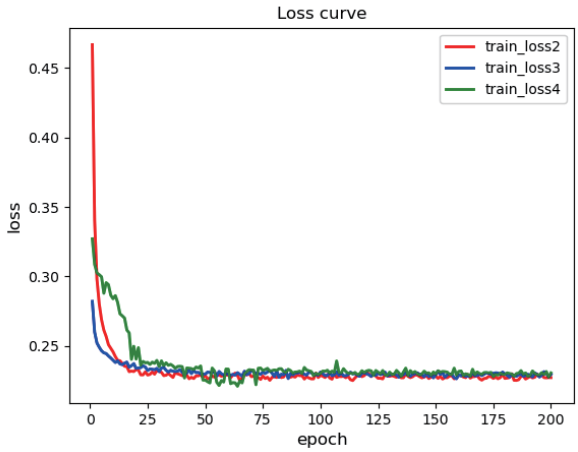

Fig. 5. Change curve of loss value during training of common dataset
Remark. Red, blue, and green curves represent change of loss value with image magnification of 2x, 3x, and 4x, respectively.

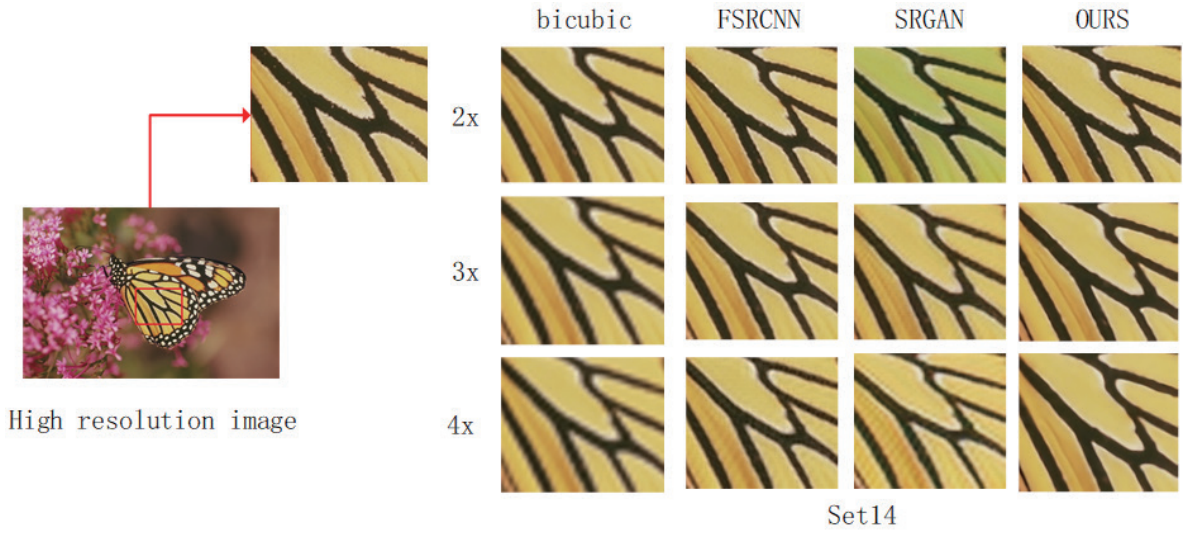

Fig. 6. Change Test results of Set14
Remark. On the left is a high-resolution image with magnified details. From left to right, each column on the right shows the test results of bicubic, FSRCNN, SRGAN, and the pro-posed method at 2x, 3x, and 4x magnifications.

Table 1. Test metric results for different methods

Dataset	Method	Magnification	PSNR	SSIM
Set14	Bicubic	2x	28.08	0.8087
		3x	27.89	0.7953
		4x	27.17	0.7180
	FSRCNN	2x	29.09	0.8203
		3x	28.03	0.8044
		4x	27.62	0.7544
	SRGAN	2x	24.29	0.8354
		3x	26.17	0.7691
		4x	23.08	0.6901
	Proposed	**2x**	**29.37**	**0.8432**
		3x	**29.88**	**0.8329**
		4x	**28.40**	**0.7769**

4.3 Results on Self-built Dataset

The loss value curves of the training process of the proposed method on the self-constructed dataset are shown in Fig. 7, where the red, blue, and green lines indicate curve changes during the training of 2x, 3x, and 4x images, respectively. These images can converge quickly during training, and the final loss values of the three magnifications from 175 to 200 epochs are maintained at around 0.6, 0.8, and 0.9, respectively, to obtain the best network model.

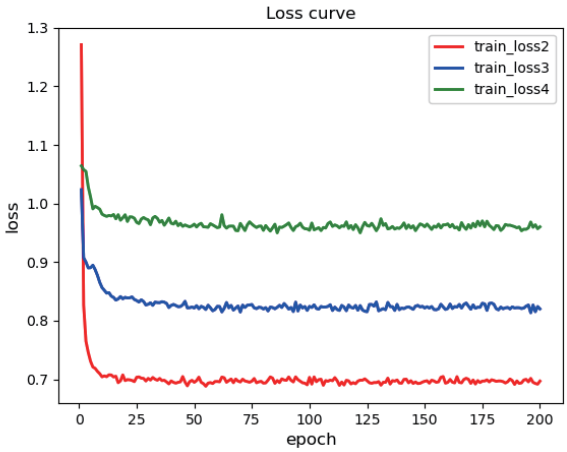

Fig. 7. Change curve of loss value during training of self-built dataset
Remark. Red, blue, and green curves represent change of loss value with image magnification of 2x, 3x, and 4x, respectively.

The test results on fish images are shown in Fig. 8 and Table 2. From Fig. 8, it can be seen that the test results of the bicubic method at 4x magnification no longer show the texture on the surface of the fish. The FSRCNN method as a whole is better than the bicubic method, with a clear texture in the middle of the image, but the details are not processed well enough at the edges, and the edge details are blurred. The test result of SRGAN shows a layer of red light at the top right, which is not in the original image, and the distortion is serious. The test results of the proposed method show clear texture and details.

It can be seen from Table 2 that the PSNR and SSIM of the proposed method reached 40.77 and 0.9777, respectively, at 2x magnification; 39.39 and 0.9667 at 3x; and 37.75 and 0.9498 at 4x. Reconstruction becomes more difficult as the magnification increases. Due to the course learning strategy introduced in this paper, which performs better at complex tasks, the image reconstruction at 4x magnification is still ideal. Therefore, the feedback network introduced in this paper can continuously optimize the shallow feature information of the low-resolution image, and the sharpness, contrast, texture, and details of the reconstructed image are closer to the original image at different magnifications. Its PSNR and SSIM, are higher than those of the bicubic, FSRCNN, and SRGAN methods. The images processed by the proposed method have the advantages of low image distortion

and high quality. Compared with other methods, the reconstructed images of the proposed method can achieve more desirable results.

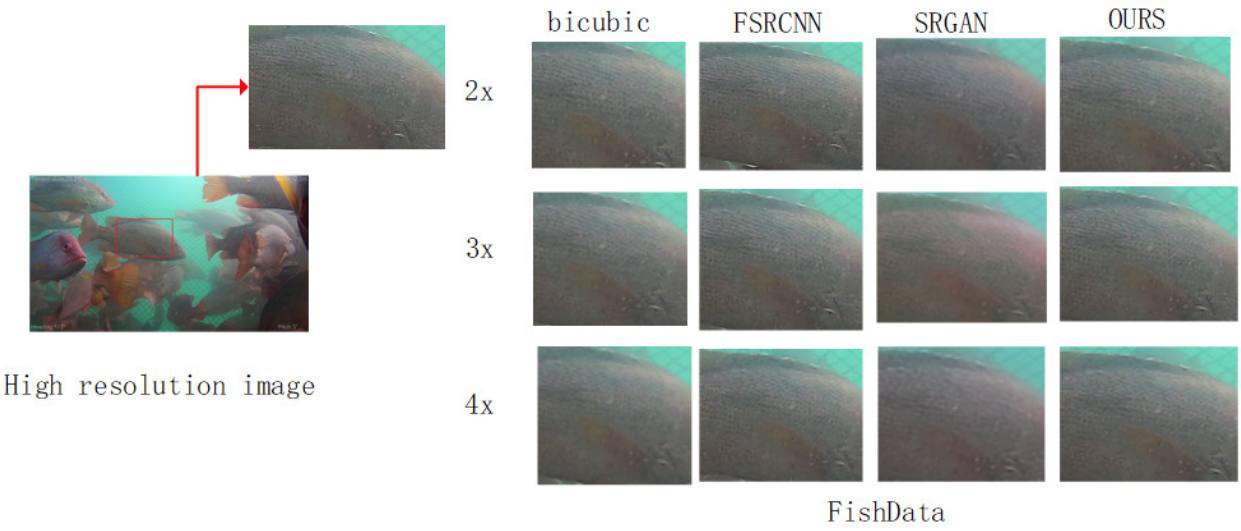

Fig. 8. Test results of self-built dataset

Table 2. Test metrics for different methods on self-built dataset

Dataset	Method	Magnification	PSNR	SSIM
Fish	Bicubic	2x	39.84	0.9609
		3x	37.34	0.9397
		4x	37.02	0.9122
	FSRCNN	2x	39.42	0.9594
		3x	38.65	0.9497
		4x	37.03	0.9176
	SRGAN	2x	29.58	0.9242
		3x	28.66	0.8437
		4x	29.39	0.8781
	Proposed	**2x**	**40.77**	**0.9777**
		3x	**39.39**	**0.9657**
		4x	**37.75**	**0.9498**

5 Conclusion

We proposed a super-resolution reconstruction method based on CNN, incorporating a feedback network implemented by a constrained RNN containing hidden states. The method was validated on the public dataset DIV2K and self-built fish dataset and compared to the bicubic, FSRCNN, and SRGAN methods. The experiments showed that the proposed method effectively alleviates the problems of blurring and unclear detail texture of reconstructed images by the other three methods, and the effect of reconstructing images is more satisfactory. PSNR and SSIM were used to evaluate the image quality, and their values were higher for the proposed method than those of the other methods for the public and fish datasets at magnifications of 2x, 3x, and 4x.

This paper makes the following contributions:

(1) We constructed a fish video image dataset with 820 images using a GLADIUS MINI robotic whale.

(2) We proposed an image super-resolution reconstruction network with feedback network. The feedback stream was used to obtain more deep information, use contextual information to better reconstruct high-resolution images, and solve the problem of unclear image texture and details. A feedback network was incorporated to process the feedback information stream, pass the deep information to the upper layer of the net-work, and refine the shallow information. A course learning strategy was introduced to enhance the applicability and robustness of the network.

The superiority of the proposed method is obvious, based on both subjective judgment by the human eye and evaluation indexes, especially on the self-constructed fish image dataset.

6 Acknowledgement

This research was funded by the Special Project on New Generation Information Technology in General Universities of Guangdong Province (2020ZDZS3008), the Special Project on Artificial Intelligence (2019KZDZS1046), and the Guangdong Rural Science and Technology Specialists Project (KTP20200219).

References

[1] T.-X. Zhai, D.-L. Zhou, C.-W. Zhang, Intelligence fishpond alarming system based on IoT, in: Proc. Internet of Things and Wireless Communication -- The 2nd National Academic Conference on Internet of Things Technology and Application and the 11th National Academic Conference on Radio Application and Management. China Institute of Communications, 2016.
[2] X.-M. Wei, Intelligent Monitoring System of Fisher Aquaculture Based on Internet of Things, [dissertation] Tianjin: Tianjin University of Technology, 2019
[3] K. Li, S.-H. Yang, R.-T. Dong, X.-Y. Wang, J.-Q. Huang, Survey of single image super-resolution reconstruction, IET Image Processing 14(11)(2020) 2273-2290.
[4] Y.-B. Zhang, Z.-H. Zhang, T.-L. Li, Research on Image Super-resolution Reconstruction Based on Deep Learning, Journal of Physics: Conference Series 1802(4)(2021) 042034.
[5] Y. Liu, L. Zhu, K.-P. Lim, Y.-H. Li, F.-P. Wang, J. Lu, Review and prospect of image super-resolution technology, Journal of Frontiers of Computer Science and Technology 14 (2)(2020) 181-199.
[6] X. Hao, G. Zhang, S. Ma, Deep Learning, International Journal of Semantic Computing 10(3)(2016) 417-439.
[7] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Networks, in: Proc. Advances in Neural Information Processing Systems, 2014.
[8] H.-P. Xie, K.-L. Xie, H.-T. Yang, Research progress of image super-resolution methods, Computer Engineering and Applications 56(19)(2020) 34-41.
[9] C. Dong, C.-C. Loy, K.-M. He, X. Tang, Learning a Deep Convolutional Network for Image Super-Resolution, in: Proc. European Conference on Computer Vision (ECCV), 2014.
[10]C. Dong, C.-C. Loy, X. Tang, Accelerating the Super-Resolution Convolutional Neural Network, in: Proc. European Conference on Computer Vision, 2016.
[11]W. Shi, J. Caballero, F. Huszár, J. Totz, A.P. Aitken, R. Bishop, D. Rueckert, Z. Wang, Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. in: Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.
[12]J. Kim, J.K. Lee, K.M. Lee, Accurate Image Super-Resolution Using Very Deep Convolutional Networks, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016.
[13]J. Kim, J.K. Lee, K.M. Lee, Deeply-Recursive Convolutional Network for Image Super-Resolution, in: Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[14]Y.-X. Guo, W. Yang, Q. Liu, Y. Wang, Survey of residual network, Application Research of Computers 37(5)(2020) 1292-1297.
[15]C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[16]T. Tong, G. Li, X. Liu, Q. Gao, Image Super-Resolution Using Dense Skip Connections, in: Proc. IEEE International Conference on Computer Vision, 2017.
[17]Q. Kuang, Single Image Super Resolution Reconstruction Algorithm Based on Deep Learning, Journal of Physics: Conference Series 1852(3)(2021) 032039.
[18]H. Zhou, L. Zhou, Q.-L. Ding, Fatigue Recognition Algorithm Based on Deep Learning, Computer Science 42(3)(2015) 191-194+200.
[19]F.P. An, Z.-W. Liu, Medical image segmentation algorithm based on feedback mechanism CNN, Contrast Media & Molecular Imaging 53(2019) 6134942.
[20]L.-H. Chen, X.-M. Yang, G. Jeon, M. Anisetti, K. Liu, A trusted medical image super-resolution method based on feedback adaptive weighted dense network, Artificial Intelligence in Medicine 106(2020) 101857.
[21]J. Schmidhuber, Deep Learning in Neural Networks: An Overview, Neural Networks 61(2015) 85-117.
[22]A. Pentina, V. Sharmanska, C.H. Lampert, Curriculum Learning of Multiple Tasks, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015.
[23]Z.-H. Chen, Research o Image Super-resolution Method Based on Deep Learning, [dissertation] Guangzhou: South China University of Technology, 2018.
[24]K.-B. Zhang, D.-N. Zhu, Z. Wang, Y.-D. Yan, Survey of Super-Resolution Images Quality Assessment, Computer Engineering and Applications 55(4)(2019) 31-40+47.
[25]W. Wang, T. Zhang, X. Wang, Survey of Single Image Super-resolution Based on Deep Learning, Journal of Chinese

Computer Systems 40(9)(2019) 1891-1896.
[26]F.-Z. Nan, Q.-L. Zeng, Y.-N. Xing, Y.-R. Qian, Single Image Super-Resolution Reconstruction based on the ResNeXt Network, Multimedia Tools and Applications 79(45-46)(2020) 34459-34470.
[27]R. Keys, Cubic convolution interpolation for digital image processing, IEEE Transactions on Acoustics, Speech, and Signal Processing 29(6)(1981) 1153-1160.